Electronic Colloquium on Computational Complexity, Report No. 148 (2013)

On the Conditional Hardness of Coloring a 4-colorable Graph with Super-Constant Number of Colors

Irit Dinur * Igor Shinkar *

Abstract

For $3 \leq q < Q$ we consider the ApproxColoring(q, Q) problem of deciding for a given graph G whether $\chi(G) \leq q$ or $\chi(G) \geq Q$. It was show in [DMR09] that the problem ApproxColoring(q, Q) is NP-hard for $q = 3, 4$ and arbitrary large constant Q under variants of the Unique Games Conjecture.

In this paper we give a tighter analysis of the reduction of [DMR09] from Unique Games to the ApproxColoring problem. We find that (under appropriate conjecture) a careful calculation of the parameters in [DMR09] implies hardness of coloring a 4-colorable graph with $\log^c(\log(n))$ colors for some constant $c > 0$. By improving the analysis of the reduction we show hardness of coloring a 4-colorable graph with $\log^c(n)$ colors for some constant $c > 0$.

The main technical contribution of the paper is a variant of the Majority is Stablest Theorem, which says that among all balanced functions in which every coordinate has $o(1)$ influence, the `Majority` function has the largest noise stability. We adapt the theorem for our applications to get a better dependency between the parameters required for the reduction.

Keywords: graph coloring, hardness of approximation, majority is stablest

*Weizmann Institute of Science. Email: `irit.dinur@weizmann.ac.il,igor.shinkar@weizmann.ac.il`.
Both authors' work supported by ISF grant 1179/09, BSF grant 2008293, and ERC starting grant 239985.

ISSN 1433-8092

1 Introduction

Graph Coloring is one of the most fundamental problems in combinatorics and computer science. A graph G on n vertices is said to be q-colorable if there is an assignment of labels $\{1, \dots, q\}$ to the vertices of G, so that every two neighboring vertices receive different colors. The chromatic number of G, denoted by $\chi(G)$, is the minimal number q such that G is q-colorable. For $q < Q$ we consider the problem $\textsc{ApproxColoring}(q, Q)$: Given a graph G, decide whether $\chi(G) \leq q$ or $\chi(G) \geq Q$. It is well known that for any constant $q \geq 3$ the problem $\textsc{ApproxColoring}(q, q+1)$ is NP-hard [Kar72]. If we consider q to be some small fixed number (e.g., 3 or 4), there is a huge gap between the values of Q for which an efficient algorithm for the problem is known, and that for which hardness results exist. For example, for $q = 3$ the best known polynomial time algorithm is due to Kawarabayashi and Thorup [KT12] who solve the problem for $Q = O(n^{0.2049})$ colors (see also [KMS98, BK97, AC06, Chl07]). On the other hand, the strongest known hardness result shows that the problem is NP-hard for $Q = 5$ (see [GK04], [KLS00]). That is, for $q = 3$ the problem is open for all $5 < Q < O(n^{0.2072})$.

Many inapproximability results are shown by a reduction from the PCP theorem [AS98, ALM+98], formulated in terms of the Label Cover problem. An instance Φ of the Label Cover problem is a bipartite graph $(V \cup W, E)$, label sets R_V and R_W, and a constraint set $\Pi = \{\pi_e \subset R_V \times R_W : e \in E\}$. The goal is to find a labeling that maximizes the fraction of satisfied constraints, i.e., the fraction of constraints that are satisfied by the labels on the relevant vertices. The value of instance Φ, denoted by $val(\Phi)$, is the fraction of satisfied constraints under such assignment. A Label Cover instance has "d-to-1 property" if the label sets are $R_V = \{1, \dots, R\}$, $R_W = \{1, \dots, dR\}$ for some $R \in \mathbb{N}$, and the constraints are projections $\pi_e : R_W \to R_V$, such that for every $a \in R_V$ there are d values $b \in R_W$ that satisfy $(a, b) \in \pi_e$. Khot [Kho02] has made following conjecture.

Conjecture 1.1 (d-to-1 Conjecture [Kho02]) *For any $\epsilon > 0$ there is $R = R(\epsilon)$ such that given a d-to-1 Label Cover instance $\Phi = (V \cup W, E, \Pi)$ with label sets $R_V = \{1, \dots, R\}$ and $R_W = \{1, \dots, dR\}$, it is NP-hard to distinguish between the case where $val(\Phi) = 1$ and the case where $val(\Phi) < \epsilon$.*

1.1 Our Result

Assuming Khot's 2-to-1 conjecture it is shown in [DMR09] that the problem of coloring a 4-colorable graph with any constant number of colors is NP-hard. We give a quantitative version of this result. Specifically, we analyze the dependency between the inapproximability factor of the 2-to-1 Label Cover problem and the number Q of colors with which it is still hard to color a 4-colorable graph. Our main result is the following theorem:

Theorem 1.2 *Suppose that given a 2-to-1 Label Cover instance Φ with n vertices and label sets of size $|R_V|, |R_W| = O(\log(n))$, there is no polynomial time algorithm that distinguishes between the case where $val(\Phi) = 1$ and the case where $val(\Phi) < \frac{1}{f(n)}$ for some $f(n)$. Then, there exists no polynomial time algorithm that colors a 4-colorable graph on N vertices with $(f(N^c))^c$ colors for some constant $c > 0$.*

For example, if $f(n) = \log^{\delta}(n)$, then there exists no poly-time algorithm that colors a 4-colorable graph on N-vertices with $\log^{\delta'}(N)$ colors for some $\delta' > 0$.

The theorem improves the dependency between the inapproximability factor of 2-to-1 Label Cover and the hardness of the graph coloring problem. For comparison, the (implicit) dependency in [DMR09] is logarithmic, i.e., the soundness of $1/f(n)$ in the Label Cover is translated into hardness of coloring a 4-colorable graph with $\Omega(\log(f(n)))$ colors.

The main technical contribution of the paper is the following theorem. It follows from a variation of the *Majority is Stablest Theorem*, which has been developed in the paper of Mossel et al. [MOO10].

Theorem 1.3 *Let q be a fixed integer and let T be a symmetric Markov operator on $[q]$ with spectral radius $\rho = \rho(T) < 1$. Then, for any $\epsilon > 0$ there exist $\delta = \epsilon^{O(1)}$ and $k = O(\log(1/\epsilon))$, where the constants in the $O()$ notation depend only on ρ and q, such that the following holds: For any two functions $f, g : [q]^n \to [0,1]$ if*

$$\mathbb{E}[f] > \epsilon \quad \mathbb{E}[g] > \epsilon \quad \text{and} \quad \langle f, T^{\otimes n} g \rangle = 0,$$

then

$$\exists i \in \{1,\ldots,n\} \quad \text{such that} \quad \mathrm{Inf}_i^{\leq k}(f) \geq \delta \quad \text{and} \quad \mathrm{Inf}_i^{\leq k}(g) \geq \delta.$$

In the analogous theorem in [DMR09, Corollary 4.12], the (implicit) dependence between δ and ϵ is exponential, that is, $\delta = \exp(-1/\epsilon)$. Our contribution is a new analysis that gives a polynomial dependence between the parameters, which in turn allows us to improve the inapproximability factor in Theorem 1.2 to be polynomial rather that logarithmic in the assumed gap of the 2-to-1 Label Cover problem.

In Section 3 we prove a variant of the Majority is Stablest Theorem with adjustments for our purposes, which allows us to conclude Theorem 1.3. For the sake of completeness, in Section 4 we present the reduction of [DMR09] from 2-to-1 Label Cover problem to the ApproxColoring problem, and work out the parameters required for Theorem 1.2.

2 Preliminaries

2.1 Functions on the q-ary hypercube

Let q be a fixed integer. Let $[q]$ denote the set $\{0,\ldots,q-1\}$. For an element $x \in [q]^n$ denote by $|x|$ the number of nonzero coordinates of x. Consider the space of real valued function with domain $[q]$, or, equivalently, a vector space $\mathbb{R}^q$ with inner product defined as

$$\langle v, w \rangle = \mathbb{E}[vw] = \frac{1}{q}\sum_{i=1}^{q} v_i w_i$$

and norm of a vector defined as

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Let $\alpha_0 = \mathbf{1}, \alpha_1, \ldots, \alpha_{q-1}$ be some orthonormal basis of $\mathbb{R}^q$. It defines naturally an orthonormal basis of $\mathbb{R}^{q^n}$ by applying the n-fold tensor product. It is easy to see that the set $\{\alpha_x = \alpha_{x_1} \otimes \alpha_{x_2} \otimes \cdots \otimes \alpha_{x_n} \in \mathbb{R}^{q^n} : x \in [q]^n\}$ is indeed an orthonormal basis of $\mathbb{R}^{q^n}$. Equivalently, we may think of α_x as a function $\alpha_x : [q]^n \to \mathbb{R}$ defined by $\alpha_x(y) = \prod_{i=1}^n \alpha_{x_i}(y_i)$. Thus, any function $f : [q]^n \to \mathbb{R}$ can be written as

$$f = \sum_{x \in [q]^n} \hat{f}(\alpha_x)\alpha_x. \tag{1}$$

Next we define the notion of influence of a variable on a function, introduced to computer science by Ben-Or and Linial in [BOL89].

Definition 2.1 *Let $f : [q]^n \to \mathbb{R}$ be a function on a q-ary hypercube. The influence of the x_i on f, is defined as*

$$\mathrm{Inf}_i(f) = \mathbb{E}_{x\setminus i}[\mathrm{Var}_{x_i}[f(x)|x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n]],$$

where $x_1, \ldots, x_n$ are uniformly distributed in $[q]$.

Some standard formulas are easily checkable using independence and orthonormality.

Proposition 2.2 *Let $f : [q]^n \to \mathbb{R}$ be as in Eq. (1). Then*

$$\mathbb{E}[f] = \hat{f}(\alpha_0), \qquad \mathbb{E}[f^2] = \sum_x \hat{f}(\alpha_x)^2,$$

$$\mathrm{Var}[f] = \sum_{|x|>0} \hat{f}(\alpha_x)^2, \qquad \mathrm{Inf}_i(f) = \sum_{x:x_i \neq 0} \hat{f}(\alpha_x)^2.$$

Analogously, we can define "low-degree influence". This notion is useful in PCPs due to the fact that a bounded function cannot have too many coordinates with non-negligible low-degree influences.

Definition 2.3 *Let $f : [q]^n \to \mathbb{R}$ be a function. Define the d-low-degree influence of f as*

$$\mathrm{Inf}_i^{\leq d}(f) = \sum_{x:x_i \neq 0, |x| \leq d} \hat{f}(\alpha_x)^2.$$

The remark above follows from the following easy proposition.

Proposition 2.4 *Let $f : [q]^n \to \mathbb{R}$ be as in Eq. (1). Then*

$$\sum_i \mathrm{Inf}_i^{\leq d}(f) \leq d \cdot \mathrm{Var}[f].$$

In particular for $f : [q]^n \to [-1, 1]$ it holds that

$$\sum_i \mathrm{Inf}_i^{\leq d}(f) \leq d,$$

and thus there are at most d/ϵ variables x_i with $\mathrm{Inf}_i^{\leq d}(f) \geq \epsilon$.

Instead of picking x at random, changing one coordinate, and seeing how it changes the value of f, we can change a constant fraction (in expectation) of the coordinates.

Definition 2.5 *Let $f : [q]^n \to \mathbb{R}$, and let $\rho \in [0, 1]$. Let $x \in [q]^n$ be chosen uniformly at random, and choose each coordinate y_i is independently to be x_i with probability ρ and uniformly random element of $[q]$ otherwise. Define the noise stability of f to be*

$$\mathbb{S}_\rho(f) = \mathbb{E}[f(x)f(y)].$$

Analogously, we generalize the notion of stability with respect to two functions:

$$\mathbb{S}_\rho(f, g) = \mathbb{E}[f(x)g(y)].$$

The notion above can be also considered as following: For any $\rho \in [0,1]$ define the following Markov operator on $[q]$ (called the Bonami-Beckner operator).

$$T_\rho = \begin{pmatrix} \rho + \frac{1-\rho}{q} & \frac{1-\rho}{q} & \dots & \frac{1-\rho}{q} \\ \vdots & \ddots & & \\ \vdots & & \ddots & \\ \frac{1-\rho}{q} & \dots & \frac{1-\rho}{q} & \rho + \frac{1-\rho}{q} \end{pmatrix}.$$

We note that $T_\rho \mathbf{1} = \mathbf{1}$ and $T_\rho v = \rho \cdot v$ for any vector $v \perp \mathbf{1}$. In particular holds $T_1(f) = f$ and $T_0(f) = \mathbb{E}[f]$. The following formulas are standard and easily checkable.

Proposition 2.6 *Let $f, g : [q]^n \to \mathbb{R}$ be as in Eq. (1) with respect to some orthonormal basis $\{\alpha_i\}$. Then*

$$T_\rho^{\otimes n}(f) = \sum_x \rho^{|x|} \hat{f}(\alpha_x)\alpha_x,$$

and by orthonormality

$$\mathbb{S}_\rho(f,g) = \langle f, T_\rho^{\otimes n} g \rangle = \sum_x \rho^{|x|} \hat{f}(\alpha_x)\hat{g}(\alpha_x).$$

By applying T_ρ on a function $f : [q]^n \to [0,1]$, the weight of f on higher levels reduces exponentially. More precisely if $g = T_\rho f$, then $\sum_{x:|x|\geq k} \hat{g}(\alpha_x)^2 \leq \rho^{2k} \sum_x \hat{f}(\alpha_x)^2 \leq \rho^{2k}$. We think of T_ρ as a smoothing operator.

Definition 2.7 *Let $g : [q]^n \to \mathbb{R}$, and let $\eta \in (0,1)$. We say that g is η-smooth if $\sum_{x:|x|\geq k} \hat{g}(\alpha_x)^2 \leq \eta^k$ for all $k \geq 0$.*

2.2 Functions in Gaussian space

Before we continue, we need to define some basic notions in $L^2(\mathbb{R}^n, \gamma)$, the space of real valued functions with domain $\mathbb{R}^n$ equipped with the standard Gaussian measure. The density function of the standard normal distribution is $\gamma(x) = \frac{1}{(2\pi)^{n/2}} \exp(-\frac{\|x\|^2}{2})$, and the inner product is defined as

$$\langle f, g \rangle = \mathbb{E}_\gamma[fg] = \int_{R^n} f(x)g(x)\gamma(x)dx.$$

For $\rho \in [-1,1]$ denote by U_ρ the Ornstein-Uhlenbeck operator

$$(U_\rho f)(x) = \mathbb{E}_{y\sim\gamma}[f(\rho x + \sqrt{1-\rho^2}y)].$$

For $\mu \in (0,1)$ define an indicator of half space function $L^2(\mathbb{R}, \gamma)$ as

$$F_\mu(x) = \mathbf{1}_{x<\Phi^{-1}(\mu)}(x),$$

where $\Phi(t) = \int_{-\infty}^t \gamma(x)dx$ is the cumulative distribution function.

A useful quantity that will appear later is $\langle F_\epsilon, U_\rho(1 - F_{1-\epsilon})\rangle$, where $\rho \in (0,1)$. Observe that $\langle F_\epsilon, U_\rho(1 - F_{1-\epsilon})\rangle = \langle F_\epsilon, U_{-\rho} F_\epsilon\rangle = \Pr[X < \Phi^{-1}(\epsilon), Y < \Phi^{-1}(\epsilon)]$, where X and Y are $(-\rho)$-correlated normal random variables with mean 0 and variance 1. That is, we have $X \sim N(0,1)$, and for independent random variable $Z \sim N(0,1)$ the random variable Y is defined to be $-\rho X + \sqrt{1-\rho^2}Z$.

It can be found in the literature (see, e.g., in [RR01], [KPW04]) that as $\epsilon \to 0$ we have

$$\langle F_\epsilon, U_{-\rho} F_\epsilon \rangle \sim \epsilon^{2/(1-\rho)} (4\pi \ln(1/\epsilon))^{\rho/(1-\rho)} \frac{(1-\rho)^{3/2}}{(1+\rho)^{1/2}}.$$

In particular if ρ is some constant bounded below 1, then

$$\langle F_\epsilon, U_\rho (1 - F_{1-\epsilon}) \rangle = \mathrm{poly}(\epsilon). \tag{2}$$

2.3 The Majority is Stablest Theorem

The Majority is Stablest Theorem [MOO10] roughly says that for all functions $f : [q]^n \to [0,1]$ in which every coordinate has $o(1)$ influence, the noise stability of f is bounded by some function of $\mathbb{E}[f]$. More specifically we have the following theorem.

Theorem 2.8 ([MOO10, Theorem 4.4]) *Fix $q \geq 2$ and $\rho \in [0,1]$. Then, for any $\epsilon > 0$ there is a small enough $\delta = \delta(\epsilon, \rho, q)$ such that for any function $f : [q]^n \to [0,1]$ that satisfies*

$$\mathrm{Inf}_i(f) \leq \delta \quad \forall i \in \{1, \ldots, n\}$$

it holds that

$$\mathbb{S}_\rho(f) \leq \left\langle F_{\mathbb{E}[f]}, U_\rho F_{\mathbb{E}[f]} \right\rangle + \epsilon.$$

In particular case of $q = 2$ and balanced functions $f : \{0,1\} \to \{0,1\}$ the theorem states that if

$$\mathbb{S}_\rho(f) > \left\langle F_{\mathbb{E}[f]}, U_\rho F_{\mathbb{E}[f]} \right\rangle + \epsilon = \frac{1}{4} + \frac{1}{2\pi} \arcsin \rho + \epsilon = \mathbb{S}_\rho(Maj) + \epsilon$$

then f has some coordinate with non-negligible influence. That is among all balanced boolean functions in which every coordinate has $o(1)$ influence, the *Majority* function has the largest noise stability.

This theorem is generalized in [DMR09] in two directions: the stability is defined with respect to two functions, and for any Markov operator T on $[q]$ (not only for T_ρ). The idea is that given a symmetric Markov operator T with eigenvalues $1 = \lambda_0 > \lambda_1 \geq \cdots \geq \lambda_{q-1}$, it is enough to bound its spectral radius $\rho = \rho(T) = \max\{|\lambda_1|, |\lambda_{q-1}|\}$ below 1. Suppose we are given a symmetric Markov operator T on $[q]$ with spectral radius $\rho < 1$, and two functions $f, g : [q]^n \to [0,1]$ that satisfy the inequality

$$\left\langle f, T^{\otimes n} g \right\rangle > \left\langle F_{\mathbb{E}[f]}, U_\rho F_{\mathbb{E}[g]} \right\rangle + \epsilon.$$

The main technical result in [DMR09, Theorem 3.1] says that in such case f and g have a common coordinate with non-negligible influence. In our setup, however, we consider functions f and g with small expectation and ρ some fixed constant. This allows us to assume $\langle f, T^{\otimes n} g \rangle > \left\langle F_{\mathbb{E}[f]}, U_{\rho'} F_{\mathbb{E}[g]} \right\rangle + \epsilon$ for some $\rho' > \rho$, and conclude that f and g have a common coordinate with relatively large influence on both functions. The exact formulation and the proof appear in the next section.

3 A Variant of the Majority is Stablest Theorem

In this section we prove our main technical result stated in Theorem 3.1, which is used in order to prove Theorem 1.3. This is a variant of [DMR09, Theorem 3.1] that we adjust for our purposes.

Let q be a fixed integer, and let T be a symmetric Markov operator on $[q]$ with eigenvalues $1 = \lambda_0 > \lambda_1 \geq \cdots \geq \lambda_{q-1} > -1$, and let $\alpha_0 = \mathbf{1}, \alpha_1, \ldots, \alpha_{q-1}$ be the corresponding eigenvectors. Denote the spectral radius of T by $\rho = \rho(T) = \max\{|\lambda_1|, |\lambda_{q-1}|\} < 1$.

Now suppose we are given two functions $f, g : [q]^n \to [0,1]$ that do not have a common influential coordinate. We show that it implies bounds on the quantity $\langle f, T^{\otimes n} g\rangle$.

Theorem 3.1 *Let q be a fixed integer, and let T be a symmetric Markov operator on $[q]$ such that $\rho = \rho(T) < 1$ and let $\rho' \in (\rho, 1)$. Then, for any $\epsilon > 0$ there are $\delta = \epsilon^{O(1)}$ and $k = O(\log(1/\epsilon))$, where the constants in the $O()$ notation depend only on $\frac{\rho}{\rho'}$ and q, such that the following holds: If $f, g : [q]^n \to [0,1]$ are two functions with $\mu = \mathbb{E}[f]$ and $\nu = \mathbb{E}[g]$ satisfying*

$$\forall i \quad \min\left(\mathrm{Inf}_i^{\leq k}(f), \mathrm{Inf}_i^{\leq k}(g)\right) < \delta,$$

then

$$\left\langle f, T^{\otimes n} g\right\rangle \geq \left\langle F_\mu, U_{\rho'}(1 - F_{1-\nu})\right\rangle_\gamma - \epsilon, \tag{3}$$

and

$$\left\langle f, T^{\otimes n} g\right\rangle \leq \left\langle F_\mu, U_{\rho'} F_\nu\right\rangle_\gamma + \epsilon. \tag{4}$$

Observe that compared to the analogous theorem of [DMR09, Theorem 3.1], we gain a better tradeoff between ϵ and δ. Specifically, we allow δ to be $\mathrm{poly}(\epsilon)$, i.e., not too small, (instead of $\delta = \exp(-1/\epsilon)$ implicitly appearing in [DMR09]). On the other hand, we get a bound on $\langle f, T^{\otimes n} g\rangle$ as a function of $\rho' \in (\rho, 1)$ rather than ρ.

For our application of the theorem, we think of ρ and ρ' as constants smaller than 1, and of μ and ν as small quantities compared to ϵ. In this setup, the polynomial dependency between ϵ and δ improves the dependency of [MOO10] and [DMR09]. The following corollary proves Theorem 1.3.

Corollary 3.2 *Let q be a fixed integer and T be a symmetric Markov operator on $[q]$ with spectral radius $\rho = \rho(T) < 1$. Then, for any $\epsilon > 0$ there exist $\delta = \epsilon^{O(1)}$ and $k = O(\log(1/\epsilon))$, where the constants in the $O()$ notation depend only on ρ and q, such that the following holds: For any two functions $f, g : [q]^n \to [0,1]$ if*

$$\mathbb{E}[f] > \epsilon \quad \mathbb{E}[g] > \epsilon \quad \textit{and} \quad \left\langle f, T^{\otimes n} g\right\rangle = 0,$$

then

$$\exists i \in \{1, \ldots, n\} \quad \textit{such that} \quad \mathrm{Inf}_i^{\leq k}(f) \geq \delta \quad \textit{and} \quad \mathrm{Inf}_i^{\leq k}(g) \geq \delta.$$

Proof Let $\rho' = \sqrt{\rho}$ (note that $\rho < \rho' < 1$), and let $\epsilon_0 = \left\langle F_\epsilon, U_{\rho'}(1 - F_{1-\epsilon})\right\rangle_\gamma$. Then

$$\left\langle f, T^{\otimes n} g\right\rangle = 0 < \left\langle F_{\mathbb{E}[f]}, U_{\rho'}(1 - F_{1-\mathbb{E}[g]})\right\rangle_\gamma - \epsilon_0.$$

By applying the contrapositive of Theorem 3.1 we get $\delta = \mathrm{poly}(\epsilon_0)$, $k = O(\log(1/\epsilon_0))$, and some $i \in \{1, \ldots, n\}$ such that $\mathrm{Inf}_i^{\leq k}(f) > \delta$ and $\mathrm{Inf}_i^{\leq k}(g) > \delta$. Using Eq. (2) we have $\epsilon_0 = \left\langle F_\epsilon, U_{\rho'}(1 - F_{1-\epsilon})\right\rangle_\gamma = \mathrm{poly}(\epsilon)$, where the degree of the polynomial depends only on ρ'. Therefore, we have $\delta = \mathrm{poly}(\epsilon)$, $k = O(\log(1/\epsilon))$, as required. ■

3.1 Proof of Theorem 3.1

Note that Eq. (3) follows from Eq. (4). Indeed, apply Eq. (4) to $1-g$ to obtain

$$\left\langle f, T^{\otimes n}(1-g)\right\rangle \leq \left\langle F_\mu, U_{\rho'} F_{1-\nu}\right\rangle_\gamma + \epsilon,$$

and then use the equalities

$$\left\langle f, T^{\otimes n}(1-g)\right\rangle = \langle f, 1\rangle - \left\langle f, T^{\otimes n} g\right\rangle = \mu - \left\langle f, T^{\otimes n} g\right\rangle = \left\langle F_\mu, U_{\rho'} 1\right\rangle - \left\langle f, T^{\otimes n} g\right\rangle.$$

So our goal in this section is to prove Eq. (4).

The following lemma is the first step in the proof of Theorem 3.1. It is proven in Lemma 3.9 in [DMR09], and essentially follows from the one function analogue proven in Theorem 4.4 in [MOO10].

Lemma 3.3 ([DMR09, Lemma 3.9]) *Let T be a symmetric Markov operator on $[q]$. Denote its eigenvalues by $1=\lambda_0 > \lambda_1 \geq \cdots \geq \lambda_{q-1} > -1$, and let $\alpha_0 = \mathbf{1}, \alpha_1, \ldots, \alpha_{q-1}$ be the corresponding orthonormal eigenbasis. Suppose that the spectral radius of T is $\rho = \rho(T) < 1$. Then, for any $\eta < 1$, $\epsilon > 0$ there is $\delta = \epsilon^{O(\frac{\log(q)}{1-\eta})}$ such that the following holds: Let $f, g : [q]^n \to [0,1]$ be two functions with $\mathbb{E}[f] = \mu$, $\mathbb{E}[g] = \nu$ with decompositions as in (1). If both functions are η-smooth, i.e.,*

$$\forall k \sum_{x:|x|\geq k} \hat{f}(\alpha_x)^2 \leq \eta^k \quad \textit{and} \quad \forall k \sum_{x:|x|\geq k} \hat{g}(\alpha_x)^2 \leq \eta^k,$$

and all influences in both of them are upper bounded by δ, that is,

$$\forall i \ \mathrm{Inf}_i(f) < \delta \qquad \textit{and} \qquad \forall i \ \mathrm{Inf}_i(g) < \delta,$$

then

$$\left\langle f, T^{\otimes n} g\right\rangle \leq \langle F_\mu, U_\rho F_\nu\rangle + \epsilon.$$

We next prove Theorem 3.1 by showing that the hypothesis of the theorem implies Eq. (4). In order to prove Theorem 3.1 let f and g be two functions as in the hypothesis of the theorem. Note that we cannot apply Lemma 3.3 directly, as we don't know that *all* variables of both f and g have small influence and both functions are smooth. We modify the functions so that they satisfy the conditions of Lemma 3.3, while making sure that the quantity $\langle f, T^{\otimes n} g\rangle$ does not change too much.

In order to satisfy the condition on influences we observe that coordinates that have large influence either on f or on g, make small contribution to $\langle f, T^{\otimes n} g\rangle$. This has been done in [DMR09].

One possible approach to satisfy the smoothness condition is to smooth f and g a little, that is to define $f_1 = T_{1-\epsilon}(f)$ and $g_1 = T_{1-\epsilon}(g)$ and show that $|\langle f, T^{\otimes n} g\rangle - \langle f_1, T^{\otimes n} g_1\rangle| < \epsilon$. But then f_1 and g_1 are only $(1-\epsilon)$-smooth, and δ that we get from Lemma 3.3 will be exponential in ϵ.[1]

A different approach is to use the fact that ρ is some constant smaller than 1 and to define $f_1 = T_\eta(f)$ and $g_1 = T_\eta(g)$ for some constant $\eta \in (\rho, 1)$. Then f_1 and g_1 are η-smooth

[1]Recall that for any $\gamma > 0$ the linear operator T_γ on $\mathbb{R}^q$ is defined by: $T_\gamma \mathbf{1} = \mathbf{1}$ and $T_\gamma v = \gamma v$ for $v \perp \mathbf{1}$. It is easy to see that the operator $S = TT_\gamma$ has the same eigenvectors as T and the corresponding eigenvalues are $1 = \lambda_0 > \lambda_1\gamma \geq \cdots \geq \lambda_{q-1}\gamma > -1$ (as long as $\gamma < 1/\rho$).

functions and $\langle f, T^{\otimes n} g\rangle = \langle f_1, S^{\otimes n} g_1\rangle$ for some operator S whose spectral radius is larger than $\rho(T)$, but still constant smaller than 1. By applying Lemma 3.3 on f_1 and g_1 with the operator S we will get $\delta = \mathrm{poly}(\epsilon)$. We now turn to the actual proof of Theorem 3.1.

Proof Set $\eta = \frac{\rho}{\rho'} < 1$ and denote $S = TT_{\frac{1}{\eta}}$. Then S has the same eigenvectors as T, largest eigenvalue 1 and $r(S) = \frac{\rho}{\eta} = \rho' < 1$. Denote

$$f_1 = T_{\sqrt{\eta}} f = \sum \hat{f}(\alpha_x)\eta^{\frac{|x|}{2}}\alpha_x \quad \text{and} \quad g_1 = T_{\sqrt{\eta}} g = \sum \hat{g}(\alpha_x)\eta^{\frac{|x|}{2}}\alpha_x.$$

Using this notation it is easy to see that we can express $\langle f, T^{\otimes n} g\rangle$ as

$$\langle f, T^{\otimes n} g\rangle = \sum_x \hat{f}(\alpha_x)\lambda_x \hat{g}(\alpha_x) = \langle f_1, S^{\otimes n} g_1\rangle. \tag{5}$$

By applying Lemma 3.3 with operator S and parameters η and $\epsilon/2$, we obtain $\delta' = \delta_{3.3}(S, \eta, \frac{\epsilon}{2})/2 = \epsilon^{O\left(\frac{1}{1-\eta}\right)} = \mathrm{poly}(\epsilon)$, where the degree of the polynomial depends only on η and q. Let $k = O(\log(1/\epsilon))$ be such that $\eta^k < \min(\delta', \epsilon/4)$, and let $\delta = (\frac{\epsilon\delta'}{8k})^2 = \mathrm{poly}(\epsilon)$. We show that these δ and k satisfy the requirements of Theorem 3.1.

Take two functions $f, g : [q]^n \to [0,1]$ such that $\forall i \min\left(\mathrm{Inf}_i^{\leq k}(f), \mathrm{Inf}_i^{\leq k}(g)\right) < \delta$. Then f_1 and g_1 are η-smooth and satisfy the same assumption. However, we cannot apply Lemma 3.3 on them with the operator S, as the requirement is that influences of all variables in both functions are small. In order overcome this obstacle we define two functions f_2 and g_2 with small influences such that the quantities $\langle f_1, S^{\otimes n} g_1\rangle$ and $\langle f_2, S^{\otimes n} g_2\rangle$ are close to each other. Specifically, let

$$B_f = \{i : \mathrm{Inf}_i^{\leq k}(f) \geq \delta'\} \qquad B_g = \{i : \mathrm{Inf}_i^{\leq k}(g) \geq \delta'\}.$$

Note that by Proposition 2.4 we have $|B_f|, |B_g| \leq k/\delta'$. Moreover, $B_f \cap B_g = \emptyset$ since by the hypothesis holds $\forall i \ \min(\mathrm{Inf}_i^{\leq k}(f), \mathrm{Inf}_i^{\leq k}(g)) < \delta < \delta'$. We define the functions $f_2, g_2 : [q]^n \to [0,1]$ to be the average over the coordinates in B_f and B_g respectively, namely

$$f_2(y) = \mathop{\mathbb{E}}_{y_i : i \in B_f}[f_1(y)] = \sum_{x : x_{B_f} = 0} \hat{f}(\alpha_x)\eta^{\frac{|x|}{2}}\alpha_x(y),$$

and

$$g_2(y) = \mathop{\mathbb{E}}_{y_i : i \in B_g}[g_1(y)] = \sum_{x : x_{B_g} = 0} \hat{g}(\alpha_x)\eta^{\frac{|x|}{2}}\alpha_x(y).$$

Clearly $\mathbb{E}[f_2] = \mathbb{E}[f] = \mu$, $\mathbb{E}[g_2] = \mathbb{E}[g] = \nu$. We have $\mathrm{Inf}_i(f_2) = 0$ for $i \in B_f$ and $\mathrm{Inf}_i(f_2) \leq \mathrm{Inf}_i(f_1) \leq \mathrm{Inf}_i^{\leq k}(f) + \eta^k < 2\delta'$ for $i \notin B_f$. Same holds for g_2. Their smoothness follows from smoothness of f_1, g_1, and we can apply Lemma 3.3 with the operator S to get

$$\langle f_2, S^{\otimes n} g_2\rangle \leq \langle F_\mu, U_{\rho'} F_\nu\rangle + \epsilon/2. \tag{6}$$

In order to complete the proof it is left to show that

$$|\langle f_1, S^{\otimes n} g_1\rangle - \langle f_2, S^{\otimes n} g_2\rangle| \leq \epsilon/2. \tag{7}$$

Here we use the assumption that no coordinate has significant influence on both functions.

$$
\begin{aligned}
|\langle f_1, S^{\otimes n} g_1\rangle - \langle f_2, S^{\otimes n} g_2\rangle| &= \left| \sum_{x: x_{B_f \cup B_g} \neq 0} \hat{f}(\alpha_x)\hat{g}(\alpha_x) \left(\prod_{i: x_i \neq 0} \lambda_{x_i} \right) \right| \\
&\leq \sum_{\substack{x:|x|\leq k \\ x: x_{B_f\cup B_g}\neq 0}} |\hat{f}(\alpha_x)\hat{g}(\alpha_x)| + \sum_{x:|x|>k} \left| \rho^{|x|} \hat{f}(\alpha_x)\hat{g}(\alpha_x) \right| \\
[\rho < \eta] \quad &\leq \sum_{i \in B_f \cup B_g} \sum_{\substack{x:|x|\leq k \\ x_i \neq 0}} |\hat{f}(\alpha_x)\hat{g}(\alpha_x)| + \eta^k \\
[\text{Cauchy Schwartz}] \quad &\leq \sum_{i \in B_f \cup B_g} \sqrt{\mathrm{Inf}_i^{\leq k}(f)} \sqrt{\mathrm{Inf}_i^{\leq k}(g)} + \eta^k \\
[i \in B_f \Rightarrow \mathrm{Inf}_i(g) < \delta] \quad &\leq (|B_f| + |B_g|)\sqrt{\delta} + \eta^k \\
\left[|B_f|, |B_g| \leq k/\delta', \eta^k \leq \epsilon/4\right] \quad &\leq \frac{2k}{\delta'}\frac{\epsilon\delta'}{8k} + \epsilon/4 \\
&= \epsilon/2.
\end{aligned}
$$

Combining Eq. (5), Eq. (6) and Eq. (7) we get the required result $\langle f, T^{\otimes n} g\rangle \leq \langle F_\mu, U_{\rho'} F_\nu\rangle + \epsilon$. This completes the proof of Theorem 3.1. ■

4 Proof of Theorem 1.2

The proof of Theorem 1.2 follows by applying the reduction of [DMR09], and using Theorem 1.3 in order to calculate the exact parameters of the reduction. Specifically we prove the following dependence in the parameters.

Theorem 4.1 *There is a reduction from 2-to-1 Label Cover problem to* ApproxColoring *problem with the following properties: Given an instance of 2-to-1 Label Cover $\Phi = (V \cup W, E, \Pi)$ with label sets of size R and $2R$, it produces a graph G' on $|W| \cdot 4^{2R}$ vertices.*

- *If $val(\Phi) = 1$, then G' is 4-colorable.*
- *If G' contains an independent set of size ϵ, then $val(\Phi) \geq \Omega\left(\frac{\epsilon\delta^2}{k^2}\right) = \mathrm{poly}(\epsilon)$, where $\delta = \epsilon^{O(1)}$ and $k = O(\log(1/\epsilon))$ as in Theorem 1.3. In other words, if $val(\Phi) \leq \frac{1}{f(n)}$, then $\chi(G) \geq f^c(n)$ for some constant $c > 0$.*

The running time of the reduction is linear in the size of the output.

Proof The reduction will use the following Markov operator T with domain $\{0,1,2,3\}^2$.

Definition 4.2 *We define a symmetric Markov operator T with domain $\{0,1,2,3\}^2$ such that $T((x_1,x_2) \leftrightarrow (y_1,y_2)) > 0$ if and only if $\{x_1,x_2\} \cap \{y_1,y_2\} = \emptyset$.*

The operator has three types of transitions, with transitions probabilities β_1, β_2, and β_3.

- *With probability β_1 we have $(x,x) \leftrightarrow (y,y)$ where $x \neq y$.*
- *With probability β_2 we have $(x,x) \leftrightarrow (y,z)$ where x, y, z are all different.*

- *With probability β_3 we have $(x, y) \leftrightarrow (z, w)$ where x, y, z, w are all different.*

For T to be a symmetric Markov operator, we need that β_1, β_2 and β_3 are non-negative and

$$3\beta_1 + 6\beta_2 = 1, \quad 2\beta_2 + 2\beta_3 = 1.$$

For example for $\beta_1 = \frac{1}{12}$, $\beta_2 = \frac{1}{8}$, and $\beta_3 = \frac{3}{8}$ we have $\rho = \rho(T) = 5/6$

The reduction: We start with a 2-to-1 Label Cover instance $\Phi = (V \cup W, E, \Pi)$. In addition, as a small technicality, we shall assume that all vertices of W have the same degree. Each $(v, w) \in E$ is associated with a constraint π_{vw} such that for each $b \in \{1, \ldots, 2R\}$ there is a unique $a \in \{1, \ldots, R\}$ such that $(a, b) \in \pi_{vw}$ (we denote $a = \pi_{vw}(b)$) and for each $a \in \{1, \ldots, R\}$ there are exactly two $b_1, b_2 \in \{1, \ldots, 2R\}$ such that $(a, b_i) \in \pi_{vw}$ (denote $(b_1, b_2) = \pi_{vw}^{-1}(a)$). We construct $G' = (V', E')$ as follows:

- Each vertex $w \in W$ is replaced by a copy of $\{0, 1, 2, 3\}^{2R}$ (denoted by $[w]$). The set of vertices in G' is $V' = \bigcup_{w \in W}[w] = W \times \{0, 1, 2, 3\}^{2R}$.
- Let T be as in Definition 4.2. For every $w_1, w_2 \in W$ that have a common neighbor $v \in V$ let π_1, π_2 be the corresponding constraints. For $x, y \in \{0, 1, 2, 3\}^{2R}$ set an edge between (w_1, x) and (w_2, y) if and only if for all $k \in \{1, \ldots, R\}$ it holds that $\{x_{i_1}, x_{j_1}\} \cap \{y_{i_2}, y_{j_2}\} = \emptyset$, where $\pi_1^{-1}(k) = \{i_1, j_1\}$ and $\pi_2^{-1}(k) = \{i_2, j_2\}$. Equivalently, we set an edge if $T(x_{\pi_1^{-1}(k)}, y_{\pi_2^{-1}(k)}) \neq 0$ for all $k \in \{1, \ldots, R\}$,

This completes the construction of G'.

Completeness: Assume there is a labeling L that satisfies all constraints of Φ. Define the coloring of G' to be $c(w, x) = x_{L(w)}$ for all $w \in W$ and $x \in \{0, 1, 2, 3\}^{2R}$. We show that this is a legal coloring of G'. Indeed, let $((w_1, x), (w_2, y)) \in E'$ be an edge of G'. Then w_1, w_2 have a common neighbor $v \in V$. Let π_1 and π_2 be the corresponding constraints, and let $k = L(v)$. Then $\pi_1(L(w_1)) = k = \pi_2(L(w_2))$, as L satisfies all the constraints.

Since $((w_1, x), (w_2, y)) \in E'$ the sets $x_{\pi_1^{-1}(k)}$ and $y_{\pi_2^{-1}(k)}$ are disjoint, and hence $c(w_1, x) \neq c(w_2, y)$ as $c(w_1, x) = x_{L(w_1)} \in x_{\pi_1^{-1}(k)}$ and $c(w_2, y) = y_{L(w_2)} \in y_{\pi_2^{-1}(k)}$.

Soundness: Suppose that G' contains an independent set $S \subseteq V'$ such that $\frac{|S|}{|V'|} \geq \epsilon$. Our goal is to show that is such case $val(\Phi) > \text{poly}(\epsilon)$. Let J be a subset of W that make a non-negligible contribution to S

$$J = \{w \in W : \frac{[w] \cap S}{[w]} > \frac{\epsilon}{2}\}.$$

By Markov inequality we have $|J| \geq \frac{\epsilon}{2}|W|$.

For each $w \in J$ let $f_w : \{0, 1, 2, 3\}^{2R} \to \{0, 1\}$ be the indicator function of S, i.e., $f_w(x) = 1$ if and only if $(w, x) \in S$. Then $\mathbb{E}[f_w] > \epsilon/2$ for each such $w \in J$. Let $\delta = \epsilon^{O(1)}$ and $k = O(\log(1/\epsilon))$ as assured by Theorem 1.3 when it is applied on the operator T from Definition 4.2 with parameter $\epsilon/2$. Next for each $w \in J$ we define a small set of labels

$$L(w) = \{i : \mathrm{Inf}_i^{\leq 2k}(f_w) > \delta/2\}.$$

By Proposition 2.4 we have $|L(w)| < \frac{4k}{\delta}$. Next we give labels to neighbors of J in Φ.

Claim 4.3 *Let $v \in N(J)$. Let $w_1, w_2 \in N(v) \cap J$, and let π_1, π_2 be the corresponding constraints. Then, there are $i \in L_{w_1}$ and $j \in L_{w_2}$ such that $\pi_1(i) = \pi_2(j)$.*

Proof Recall that f_w's are indicators of an independent set. Thus $f_{w_1}(x) = 1 = f_{w_2}(y)$ implies that $((w_1, x), (w_2, y)) \notin E'$. Therefore, $T\left(x_{\pi_1^{-1}(k)}, y_{\pi_2^{-1}(k)}\right) = 0$ for some $k \in \{1, \dots, R\}$, and thus

$$T^{\otimes R}\left((x_{\pi_1^{-1}(1)}, \dots, x_{\pi_1^{-1}(R)}), (y_{\pi_2^{-1}(1)}, \dots, y_{\pi_2^{-1}(R)})\right) = 0.$$

Define

$$\overline{f}(x_{\pi_1^{-1}(1)}, \dots, x_{\pi_1^{-1}(R)}) = f_{w_1}(x_1, \dots, x_{2R})$$

and

$$\overline{g}(y_{\pi_2^{-1}(1)}, \dots, y_{\pi_2^{-1}(R)}) = f_{w_2}(y_1, \dots, y_{2R}),$$

where we think of $\overline{f}, \overline{g}$ as functions in R variables, each taking values in $\{0,1,2,3\}^2$. We show that $\langle \overline{f}, T^{\otimes R}\overline{g} \rangle = 0$. Then, using Theorem 1.3, we conclude that there is $\ell \in \{1, \dots, R\}$ such that $\mathrm{Inf}_\ell^{\le k}(\overline{f}) > \delta$ and $\mathrm{Inf}_\ell^{\le k}(\overline{g}) > \delta$. Using the relation between $\overline{f}$ and f_{w_1} we conclude that there is some $i \in \pi_1^{-1}(\ell)$ such that $\mathrm{Inf}_i^{\le 2k}(f_{w_1}) > \delta/2$. Similarly for g there is some $j \in \pi_2^{-1}(\ell)$ such that $\mathrm{Inf}_j^{\le 2k}(f_{w_2}) > \delta/2$. Therefore, there are $i \in L_{w_1}, j \in L_{w_2}$ such that $\pi_1(i) = \pi_2(j)$.

It is left to show that $\langle \overline{f}, T^{\otimes R}\overline{g} \rangle = 0$. Indeed,

$$\begin{aligned}
\langle \overline{f}, T^{\otimes R}\overline{g} \rangle &= \frac{1}{4^{2R}} \sum_{x \in (\{0,1,2,3\}^2)^R} \overline{f}(x) T^{\otimes R}\overline{g}(x) \\
&= \frac{1}{4^{2R}} \sum_x \overline{f}(x) \sum_y T^{\otimes R}(x,y)\overline{g}(y) \\
&= \frac{1}{4^{2R}} \sum_{\substack{x: f_{w_1}(x)=1 \\ y: f_{w_2}(y)=1}} T^{\otimes R}(x_{\pi_1^{-1}}, y_{\pi_2^{-1}}) \\
&= \frac{1}{4^{2R}} \sum_{\substack{x: f_{w_1}(x)=1 \\ y: f_{w_2}(y)=1}} 0 \\
&= 0,
\end{aligned}$$

as required. ■

From the claim above we get that for all $v \in N(J)$ and any $w_1, w_2 \in N(v) \cap J$ it holds that

$$\Pr_{\substack{i \in L(w_1) \\ j \in L(w_2)}}[\pi_1(i) = \pi_2(j)] \ge \frac{1}{|L(w_1)||L(w_2)|} \ge \left(\frac{\delta}{4k}\right)^2.$$

By averaging there is $L_0 : V \cup W \to \{1, \dots, 2R\}$ such that for all $w \in J$ we have

$$\Pr_{v \in N(w)}[L_0(v) = \pi(L_0(w))] \ge \left(\frac{\delta}{4k}\right)^2.$$

Hence, if we assume regularity on the vertices of W, we get

$$\begin{aligned}\Pr_{(v,w)\in E}[L_0(v) = \pi(L_0(w))] &\geq \Pr_{w\in W}[w \in J] \Pr_{v\in N(w)}[L_0(v) = \pi(L_0(w))|w \in J] \\ &\geq \epsilon\left(\frac{\delta}{4k}\right)^2 \\ &= \mathrm{poly}(\epsilon).\end{aligned}$$

We conclude that $val(\Phi) > \mathrm{poly}(\epsilon)$, which completes the soundness analysis of the reduction. Theorem 1.2 follows. ■

Acknowledgments

We are thankful to Oded Regev for useful comments on the paper.

References

[AC06] S. Arora and E. Chlamtac. New approximation guarantee for chromatic number. In *STOC '06: Proceedings of the thirty-eighth annual ACM Symposium on Theory of Computing*, pages 215–224, 2006.

[ALM+98] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998.

[AS98] S. Arora and S. Safra. Probabilistic checking of proofs : a new characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998.

[BK97] A. Blum and D. Karger. An $\widetilde{O}(n^{3/14})$-coloring algorithm for 3-colorable graphs. *Information Processing Letters*, 61:49–53, 1997.

[BOL89] M. Ben-Or and N. Linial. Collective coin flipping. *Randomness and Computation*, pages 91–115, 1989.

[Chl07] E. Chlamtac. Approximation algorithms using hierarchies of semidefinite programming relaxations. In *FOCS '07: Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 691–701, 2007.

[DMR09] I. Dinur, E. Mossel, and O. Regev. Conditional hardness for approximate coloring. *SIAM Journal on Computing*, 39(3):843–873, 2009.

[GK04] V. Guruswami and S. Khanna. On the hardness of 4-coloring a 3-colorable graph. *SIAM J. Discret. Math.*, 2004.

[Kar72] R. Karp. Reducibility among combinatorial problems. *Complexity of Computer Computations*, 43:85–103, 1972.

[Kho02] S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 767–775, 2002.

[KLS00] S. Khanna, N. Linial, and S. Safra. On the hardness of approximating the chromatic number. *Combinatorica*, 20(3):393–415, 2000.

[KMS98] D. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidefinite programming. *J. ACM*, 45(2):246–265, 1998.

[KPW04] E. De Klerk, D.V. Pasechnik, and J.P. Warners. On approximate graph colouring and max-k-cut algorithms based on the ϑ-function. *Journal of Combinatorial Optimization*, 8(3):267–294, 2004.

[KT12] K. Kawarabayashi and M. Thorup. Combinatorial coloring of 3-colorable graphs. *FOCS '12: Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science*, pages 68–75, 2012.

[MOO10] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Annals of Mathematics*, 171(1):295–341, 2010.

[RR01] Y. Rinott and V. Rotar'. A remark on quadrant normal probabilities in high dimensions. *Statistics and Probability Letters*, 51(1):47–51, 2001.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de